School-Wide Screening

Hugh W. Catts, Ph.D.
University of Kansas

National SEA Conference on SLD Determination
Kansas City, MO
April 19-21, 2006

Early Identification

Early Identification

↓

Early Intervention

↓

Reduced Risk for RD

3

Early Intervention

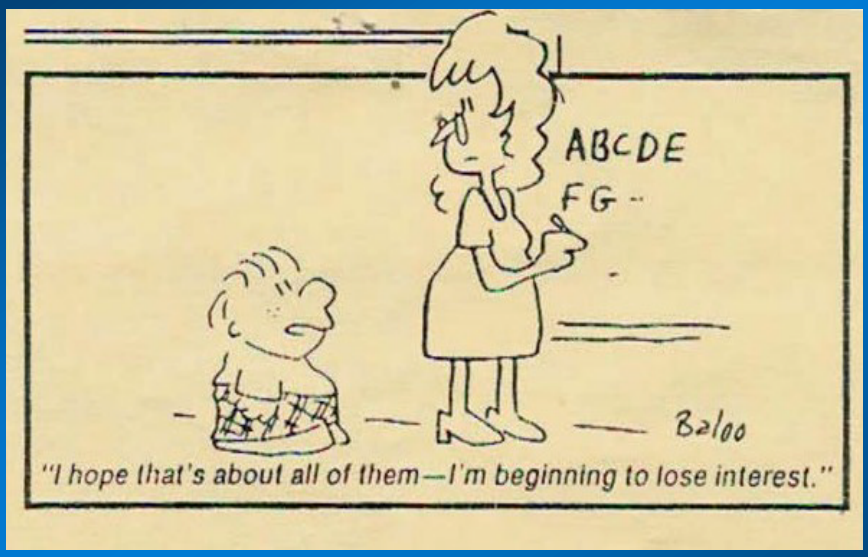
- Critical because children who start out as poor readers generally continue to be poor readers
- Poor reading achievement quickly leads to a host of negative consequences

4

Negative Consequences

- Low motivation
- Negative expectations
- Limited practice
- Academic failure

5



6

Negative Consequences

- Low motivation
- Negative expectations
- Limited practice
- Academic failure

7

School-Wide Screening

- Importance of accuracy
- What to measure
- Current screening tools
- New directions in research
- Conclusions

8

Screening

- Screening tests have a long tradition in health professions
- Used to detect potential health problems in an individual who doesn't show symptoms
- Once identified, follow-up testing is conducted, and if required, intervention is initiated to prevent or limit the condition or disease
- Common screening tests include tests for high cholesterol, early signs of cancer, depression, or hearing problems

9

Screening in Schools

- Screening tests also have a long tradition in education.
- Typically administered in kindergarten or first grade with the purpose of identifying children at risk for academic problems
- Screening takes on a more prominent role in a RTI framework

10

Screening Accuracy

- Particular attention is given to the accuracy of screening instruments
- Errors in identification can be costly
  - over identification
  - under identification

11

Public Health

- Over identification
  - expense of additional testing
  - unnecessary worry
- Under identification
  - miss serious health problem

12

Education

- Over identification
  - expense of additional testing
  - expense of early intervention services
- Under identification
  - miss opportunity for early intervention

13

Clinical Decision Making Model

14

Clinical Decision Making Model

Screen

Outcome	At risk	Not at risk
RD	True Positive	False Negative
Normal	False Positive	True Negative

15

Clinical Decision Making Model

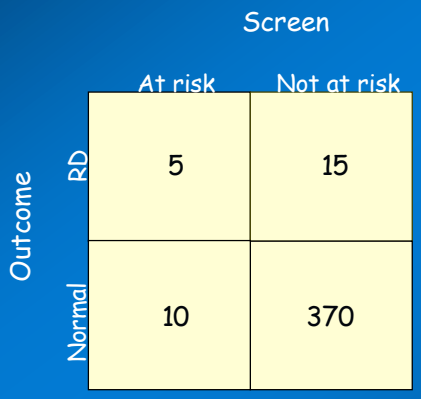
Screen

Outcome	At risk	Not at risk
RD	True Positive a	
Normal		True Negative d

Total % Correct
a+d/ total

16

Screening Accuracy

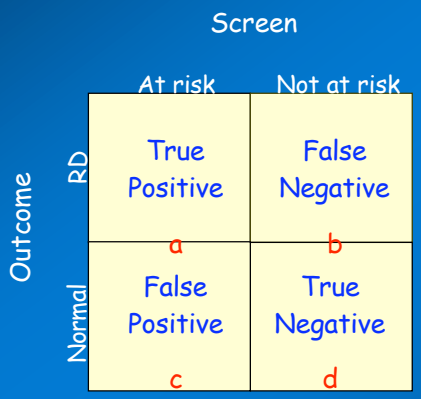


Base rate
5%

Total % Correct
94%

17

Clinical Decision Making Model

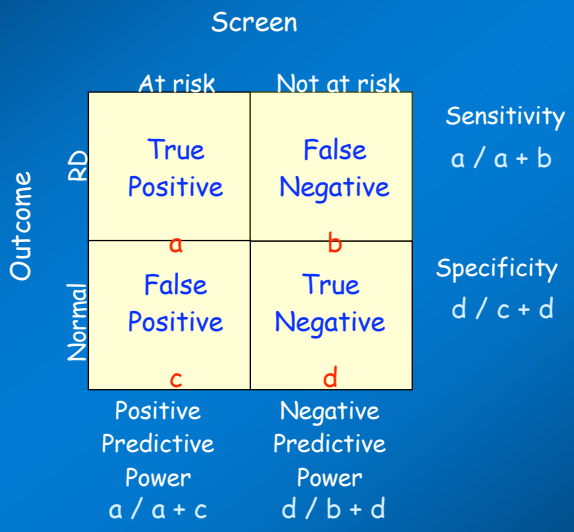


Sensitivity
a / a + b

Specificity
d / c + d

18

Clinical Decision Making Model



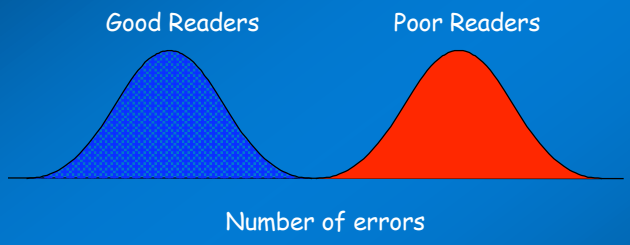
19

Accuracy of Screening is determined by …

- How well your instrument separates those who eventually will have a problem from those who will not
- What you choose as a cut-off score

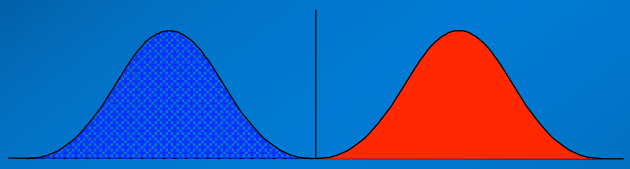
20

The Ultimate Screen



21

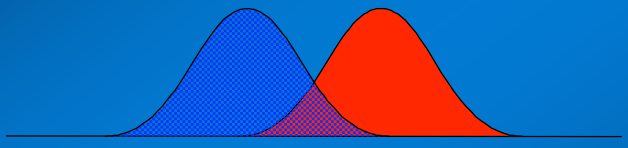
The Ultimate Screen



TP 100	FN 0
FP 0	TN 100

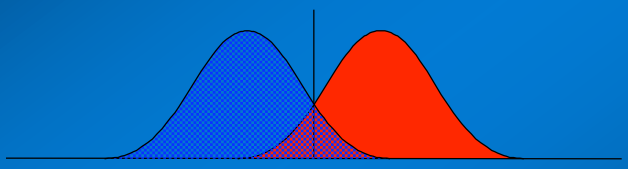
22

More Typical Screen



23

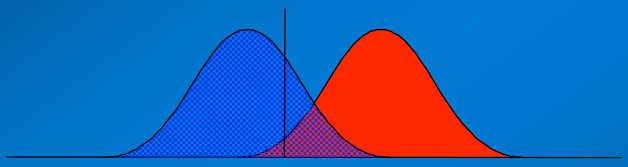
More Typical Screen



TP 80	FN 20
FP 20	TN 80

24

More Typical Screen

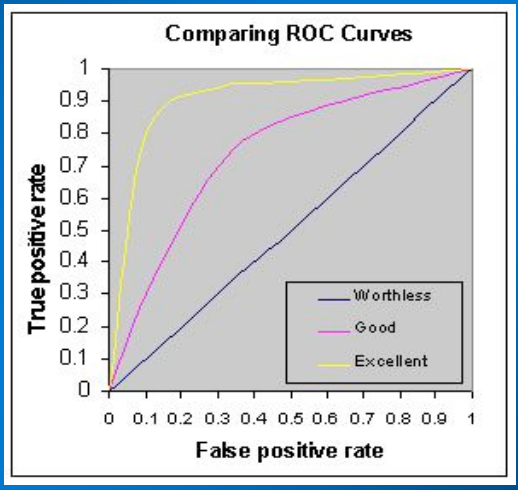


TP 95	FN 5
FP 35	TN 65

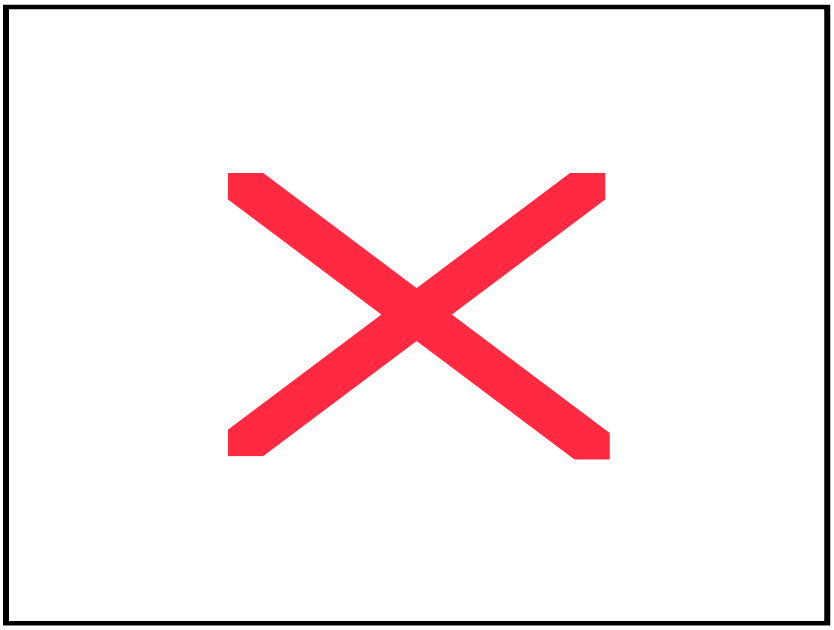
25

ROC Curve

http://www.anaesthetist.com/mnm/stats/roc/



26

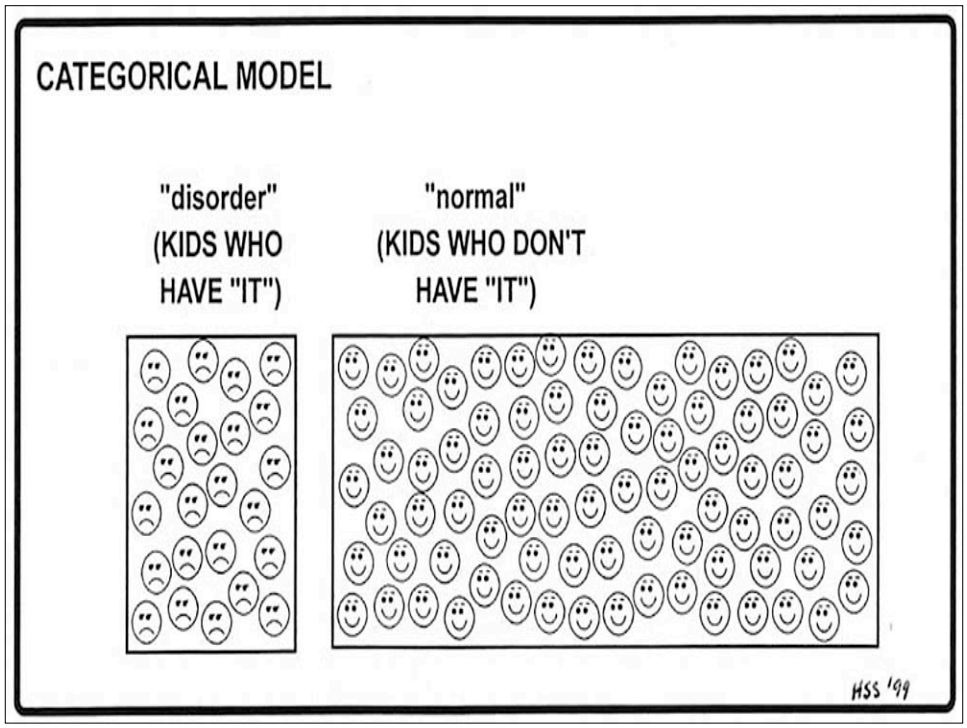


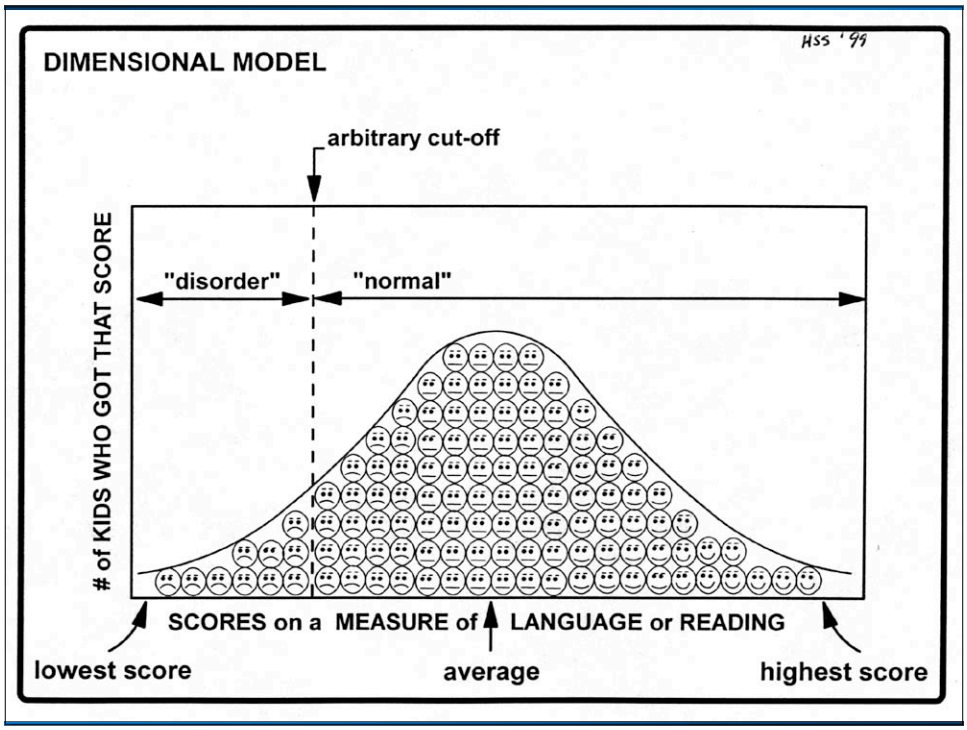
27

Base Rate

- The proportion of population that has the condition
- Sometimes base rate is straightforward -children with severe/multiple handicaps
- Base rate of RD is not clear-cut
- Reading achievement is continuously distributed with no clear demarcation between good and poor readers
- Doesn't follow a categorical model

28


CATEGORICAL MODEL
"disorder"
(KIDS WHO
HAVE "IT")
"normal"
(KIDS WHO DON'T
HAVE "IT")
HSS '99


DIMENSIONAL MODEL
HSS '99
arbitrary cut-off
"disorder"
"normal"
# of KIDS WHO GOT THAT SCORE
SCORES on a MEASURE of LANGUAGE or READING
lowest score
average
highest score

Base Rate

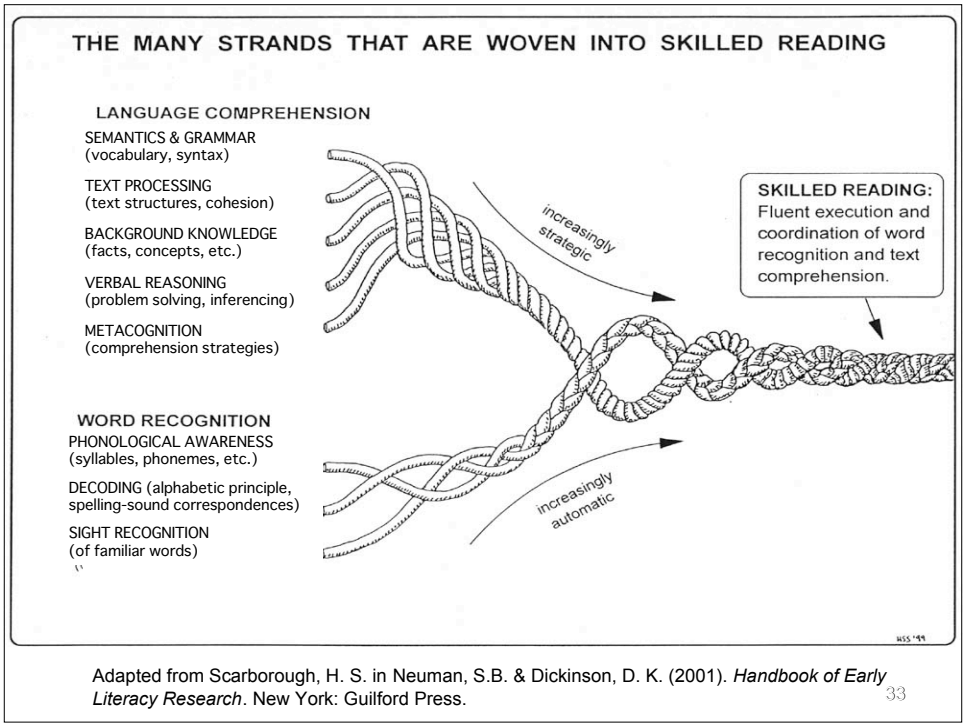
- Determined in part by perspective on the purpose of early identification
  - Traditional LD model
    - base rate 5%
    - at-risk rate 15-20%
  - Prevention-oriented general ed model
    - base rate 20-30% at-risk rate 50%
- Dependent on resources

31

What to Measure?

- What is the criterion? What are we predicting to?
- Reading comprehension
- Reading comprehension involves a mixture of complex abilities
- Role of each changes over time

32



Adapted from Scarborough, H. S. in Neuman, S.B. & Dickinson, D. K. (2001). *Handbook of Early Literacy Research*. New York: Guilford Press.

33

Predicting Comprehension

- word reading
  - letter knowledge
  - phonological awareness
- oral reading fluency
- vocabulary and grammar
- listening comprehension

34

Measures

- Need to be matched to abilities of children
- Should be consistent with the expectations of the curriculum
- Estimate of risk is a "moving target"

35

Measures

- Need to use multiple measures
- Most early predictors are only moderately correlated with reading
- Need a combination to attain high classification accuracy
- Measure more than once

36

Screening Tools

- Readily available
- Standardized
- Easily administered
- Accurate

37

Phonological Awareness Literacy Screening (PALS-K; Invernizzi, Juel, Swank, & Meier)

- http://pals.virginia.edu
- Measures kindergarten students' literacy development with the following subtests
  - Rhyme Awareness (group then individual if needed)
  - Beginning Sound Awareness (group then individual if needed)
  - Alphabet Knowledge
  - Letter Sounds
  - Spelling (group then individual if needed)
  - Concept of Word
- Takes approximately 30-45 minutes to complete
- A summed score is obtained which can be used to compare to benchmarks (fall and spring)
- PALS-PreK and PALS 1-3 also available
- Classification accuracy of combined PALS K, 2-3 to state assessment was 82%

38

Texas Primary Reading Inventory (Foorman et al., 1998- www.tpri.org)

- Designed to be used by teachers to identify children at risk for RD and to further evaluate their strengths and weaknesses in reading-related skills
- 5 screens for K-2nd grade
- Designed to hold false negatives at a minimum
- Includes an inventory of secondary measures to help rule out false positives and inform instruction

39

TPRI (1998)

K (Dec) predicting end of 1st

Screen (shorten version)

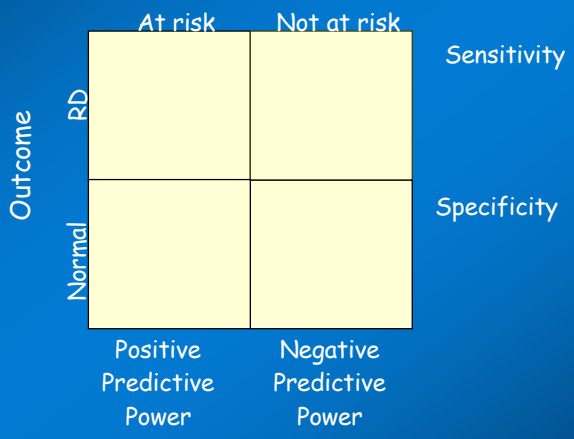
- Letter-sound identification (10 items)
- Phoneme blending (8 items)

40

TPRI (1998)

K (Dec) predicting end of 1st

Screen (shorten version)

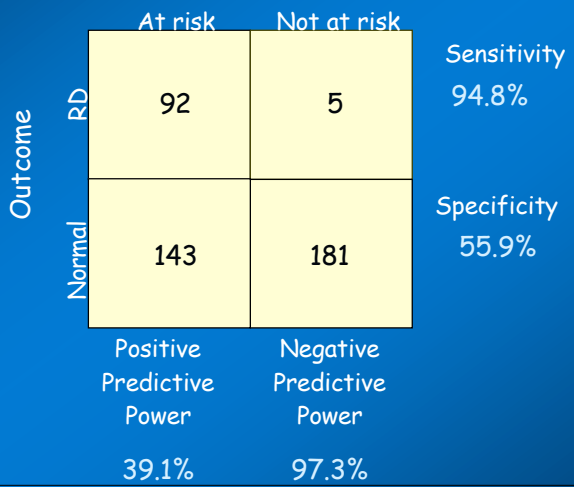


41

TPRI (1998)

K (Dec) predicting end of 1st

Screen (shorten version)



42

Texas Primary Reading Inventory (Foorman et al., 1998- www.tpri.org)

- Inventory of secondary measures (12 measures)
  - book and print awareness
  - rhyming
  - blending word parts
  - blending phonemes
  - deleting initial sounds
  - deleting final sounds
  - letter-name identification
  - letter to sound linking A & B
  - listening comprehension 1-3
- Most have 5 items
- Designed to progress for easy to difficult
- About 20 minutes to administer

43

Dynamic Indicators of Basic Early Literacy Skills (DIBELS)

- Standardized and readily available

  www.dibels.uoregon.edu

  www.aimsweb.com
- Curriculum-Based Measurement Tool (CBM)
- Developed to monitor progress and inform instruction

44

CBM Tools

- Short assessments
- Most often involve speeded performance
- Multiple forms
- Tied to curriculum

45

CBM Tools

- Letter-Name Fluency
- Letter-Sound Fluency
- Initial-Sound Fluency
- Phoneme Segmentation Fluency
- Nonword Reading Fluency
- Word Identification Fluency
- Oral Reading Fluency
- Oral Retell Fluency
- Maze Fluency

46

CBM Tools

- Assessments given 3 or more times a year to evaluate growth in reading (meeting benchmarks)
- Each can be considered a screening opportunity

47

DIBELS

K (Fall) predicting end of 1st

Screen (Initial sound fluency, Letter name fluency)

48

DIBELS

K (Fall) predicting end of 1st*

Screen (Initial sound fluency, Letter name fluency)

Outcome	At risk	Not at risk
RD	8577	1824
Normal	9345	12258

Sensitivity 82.5%

Specificity 56.7%

Base rate 32.5%

Risk rate 56.0%

Positive Predictive Power 47.9%

Negative Predictive Power 87.0%

*Adapted from summary data with important qualifications

49

First Grade Screening

50

TPRI (1998)

1st (Oct) predicting end of 1st

Screen (letter-sound, blending, word reading)

Outcome	At risk	Not at risk	
RD	111	8	Sensitivity 93.3%
Normal	175	305	Specificity 63.5%
	Positive Predictive Power 38.8%	Negative Predictive Power 97.4%	

Base rate 19.9%

Risk rate 47.7%

51

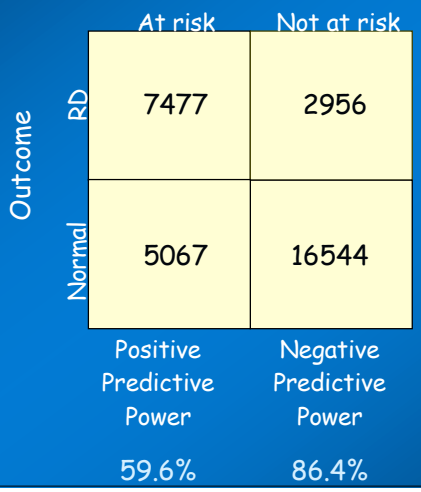
DIBELS

1st NWF predicting end of 1st ORF

52

DIBELS

1st NWF predicting end of 1st ORF



Sensitivity 71.7%

Specificity 76.6%

Base rate 32.6%

Risk rate 39.1%

*Adapted from summary data with important qualifications

53

Dynamic Assessment

- Measurement of ability over time in order to monitor progress
- Measurement of learners' potential over the short term
- Assessor actively intervenes during the course of the assessment with the goal of intentionally inducing changes in the learner's current level of performance.
- "Mini-assessment" of response to intervention

54

O'Connor & Jenkins (1999)

Oct 1st predicting April 1st

Screen (phoneme seg, RLN, phoneme repetition)

Outcome	At risk	Not at risk
RD		
Normal		

Base rate

5.1%

55

O'Connor & Jenkins (1999)

Oct 1st predicting April 1st

Screen (phoneme seg, RLN, phoneme repetition)

Outcome	At risk	Not at risk
RD	11	0
Normal	26	178

Sensitivity

100%

Specificity

87.3%

Base rate

5.1%

Risk rate

17.2%

Positive Predictive Power

29.7%

Negative Predictive Power

100%

56

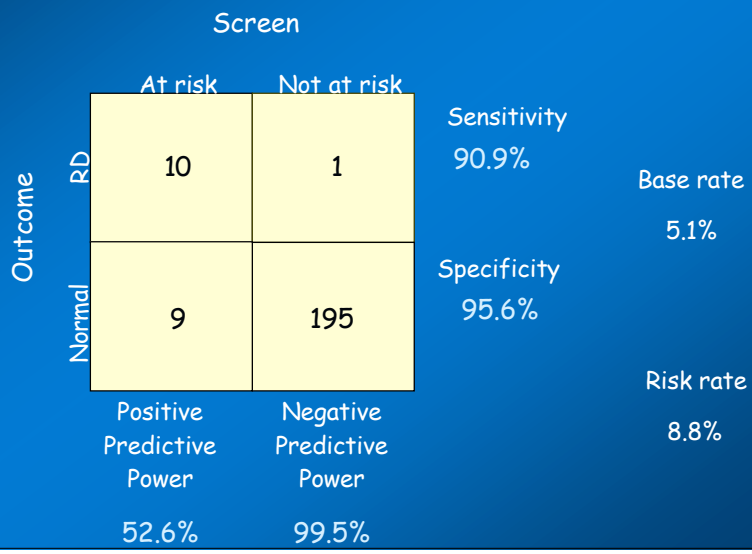
O'Connor & Jenkins (1999)

- Dynamic Assessment
  - taught at-risk children phoneme segmentation using a set of test items
  - score based on the number of trials needed to master the task

57

O'Connor & Jenkins (1999)

Oct 1st predicting April 1st (dynamic)



58

Compton, Fuchs, Fuchs, & Bryant (in press)

- Screened in 1st (Oct) predicting end of 2nd
- Measures
  - CTOPP Sound Matching
  - CTOPP Rapid Digit Naming
  - WJPB-R Oral Vocabulary
  - Word Identification Fluency (WIF)

    Initial level, 5-week slope

59

Grade 1 Word-Identification Fluency

Teacher: ***Read these words.***

Time: 1 minute.

two
for
come
because
last
from
...

60

Compton, Fuchs, Fuchs, & Bryant (in press)

- Screened in 1st (Oct) predicting end of 2nd
- Measures
  - CTOPP Sound Matching
  - CTOPP Rapid Digit Naming
  - WJPB-R Oral Vocabulary
  - Word Identification Fluency (WIF)

    Initial level, 5-week slope

61

Compton et al. (in press)

1st (Oct) predicting end of 2nd

Screen (includes WIF level & slope)

Outcome	At risk	Not at risk	
RD	35	2	Sensitivity 94.6%
Normal	14	155	Specificity 91.7%
	Positive Predictive Power 71.4%	Negative Predictive Power 98.7%	

62

Beyond First grade

- Most common screening for Tier 2 has been measure of ORF
- ORF strongly correlated with 3rd grade state assessments
- High correlations do not necessarily translate into high sensitivity and specificity

63

Concurrent Validity

	R				
OSA (Good, Simmons, & Kame'enui, 2001)	.73				
FCAT-SSS (Buck & Torgesen, 2003)	.70				
ISAT (Sibley, Biwer, & Hesch, 2001)	.79				
ASA (Linner, 2001)	NA				
CSAP (Shaw & Shaw, 2002)	.80				
MEAP (4th grade) (McGlinchey & Hixson, 2004)	.49-.81				

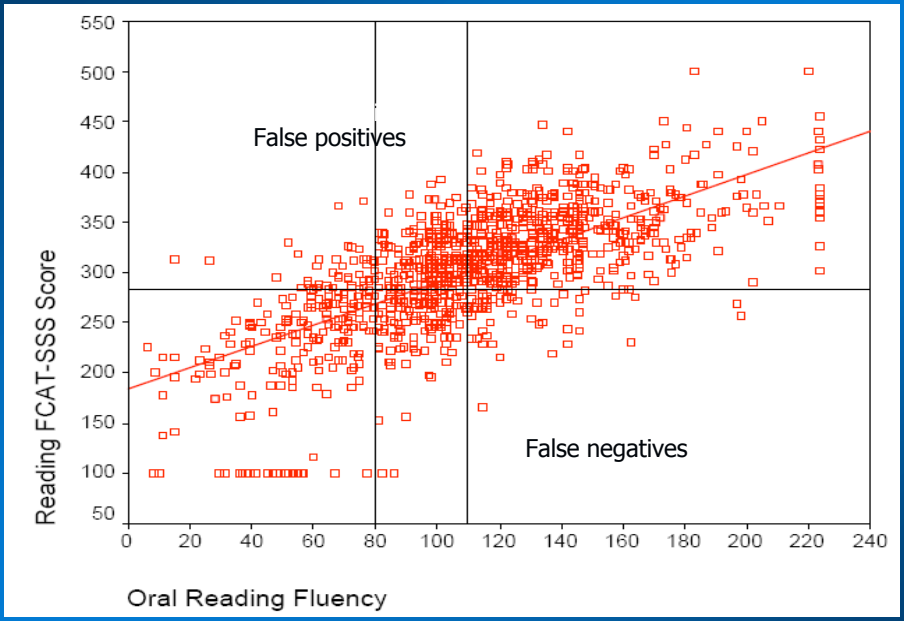
Concurrent Validity

	R	Sensitivity	Specificity	Positive Predictive Power	Negative Predictive Power
OSA (Good, Simmons, & Kame'enui, 2001)	.73	89.4	71.3	43.7	96.4
FCAT-SSS (Buck & Torgesen, 2003)	.70	85.3	69.0	57.3	90.6
ISAT (Sibley, Biwer, & Hesch, 2001)	.79	93.8	74.5	37.5	98.6
ASA (Linner, 2001)	NA	89.7	74.3	44.3	96.9
CSAP (Shaw & Shaw, 2002)	.80	80.0	62.8	42.9	90.0
MEAP (4th grade) (McGlinchey & Hixson, 2004)	.49-.81	75.0	74.0	77.0	72.0

CBM & State Assessments

- Reported results are usually much better
- Most reports only consider the low and high risk groups
- Students in the "some risk" category are not included
- Equally likely to have good vs. poor outcomes
- But results should be expected on the basis of the simple view

66



Buck & Torgesen (2003)

67

Possible Solutions

- Measurement of level and slope may help (e.g., dual discrepancy)
- Add assessments of language abilities

68

What have we learned about screening?

- Can identify children at risk for reading problems
- Can be done as early as the fall of kindergarten
- Need to choose measures carefully
- Must match measures to curriculum
  - letter naming
  - phonological awareness
  - word reading
  - text reading
- Must not forget about other factors related to comprehension
  - oral language

69

What have we learned about screening?

- False positive rates are high and efforts need to be in place to limit the cost of over prediction
- Brief secondary assessments (TPRI)
- Duel discrepancy
- Short-term instruction (dynamic assessment)
- Tier 2 (RTI)

70